

PCI Express® Link Training and Protocol Debug Techniques

Gordon Getty

Senior Applications Engineer

Teledyne LeCroy

Copyright © 2017 PCI-SIG® - All Rights Reserved

Disclaimer

<u>Presentation Disclaimer</u>: All opinions, judgments, recommendations, etc. that are presented herein are the opinions of the presenter of the material and do not necessarily reflect the opinions of the PCI-SIG®.

Copyright © 2017 PCI-SIG® - All Rights Reserved

Objective

- **The examples are all real problems that have been observed multiple times.**
- **This presentation will discuss techniques applicable to testing Link Training and Protocol for all PCI Express® devices, including those designed to the PCI Express 4.0 specification.**

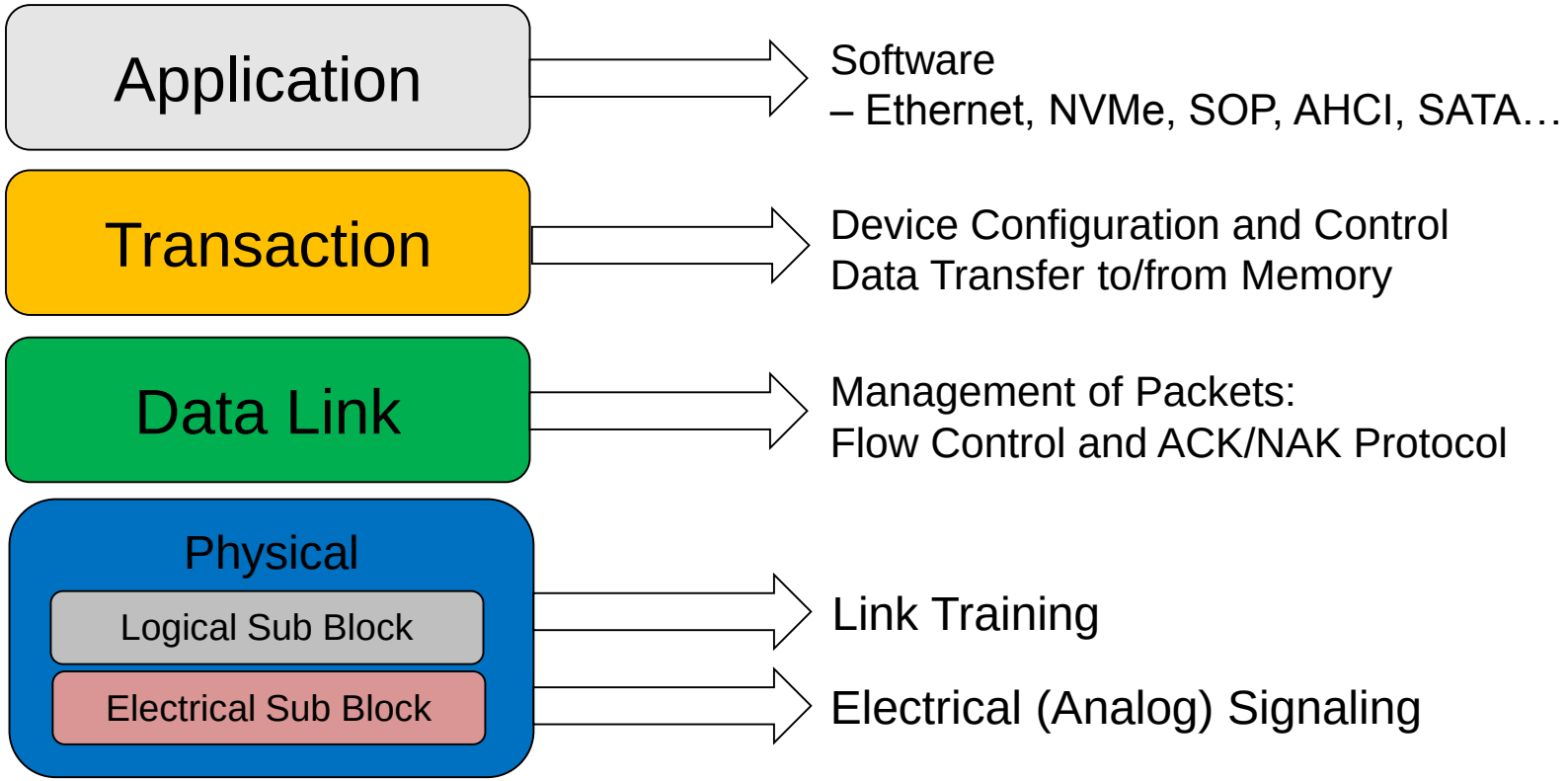
Copyright © 2017 PCI-SIG® - All Rights Reserved

Agenda

- **PCI Express Layered Model Overview**
- **Top PCIe® Protocol Problems**
  - Link Training
  - Polarity
  - Lane Reversal
  - Power Management
  - Reset Issues
  - Flow Control Problems
  - Class Code
  - Compliance Test Failures
  - Device ID Changes

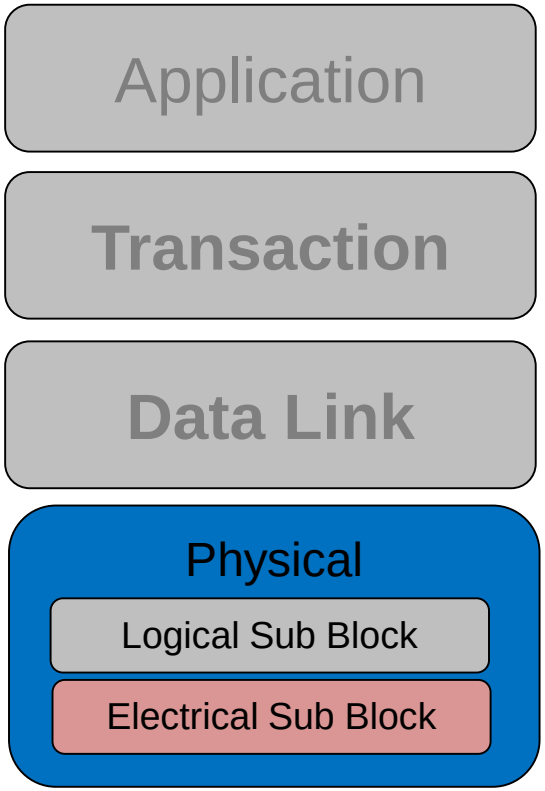
Copyright © 2017 PCI-SIG® - All Rights Reserved

PCI Express Layered Model



Copyright © 2017 PCI-SIG® - All Rights Reserved

Link Training Problems



Copyright © 2017 PCI-SIG® - All Rights Reserved

My Device Does Not Show in Device Manager

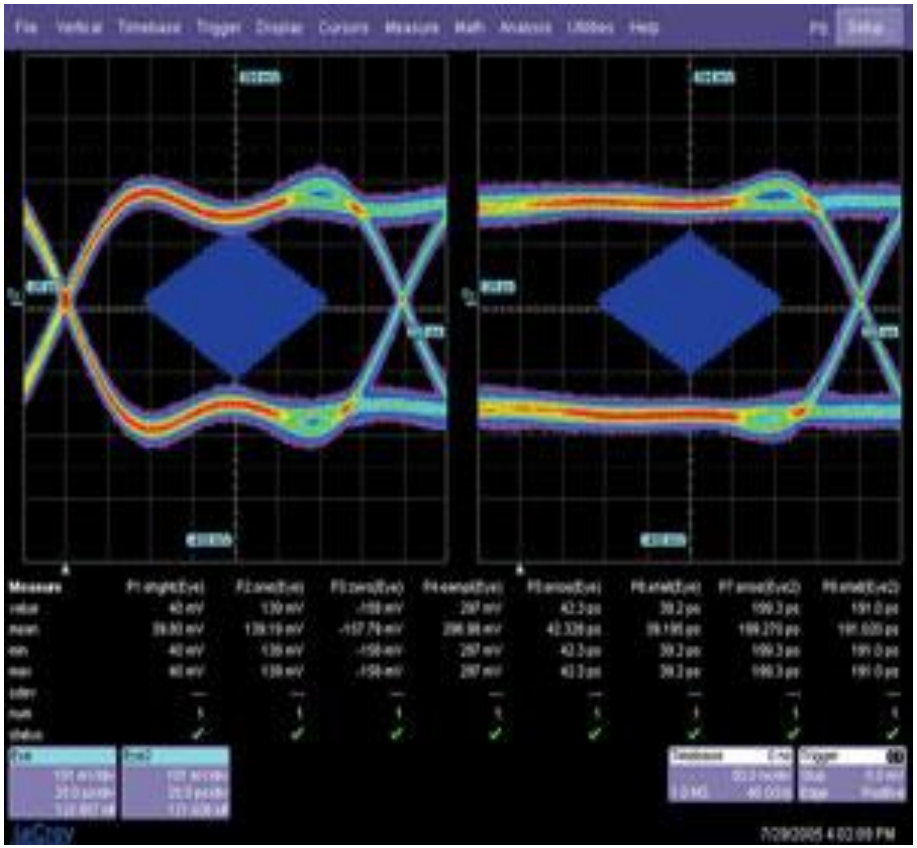
- **Is the Link Active?**
- **Do a snapshot trigger**
  - If the link is active, you will see Update_FC packets and SKP ordered sets
  - If the link is not active, you may see TS1 or TS2 ordered sets
  - If the link is unstable, you may see Update_FC and SKP ordered sets, with occasional TS1/TS2 sequences as the link goes to Recovery
- **Bottom Line: If the link is not active, there is no way the device manager will see the device.**

Copyright © 2017 PCI-SIG® - All Rights Reserved

Signal Integrity

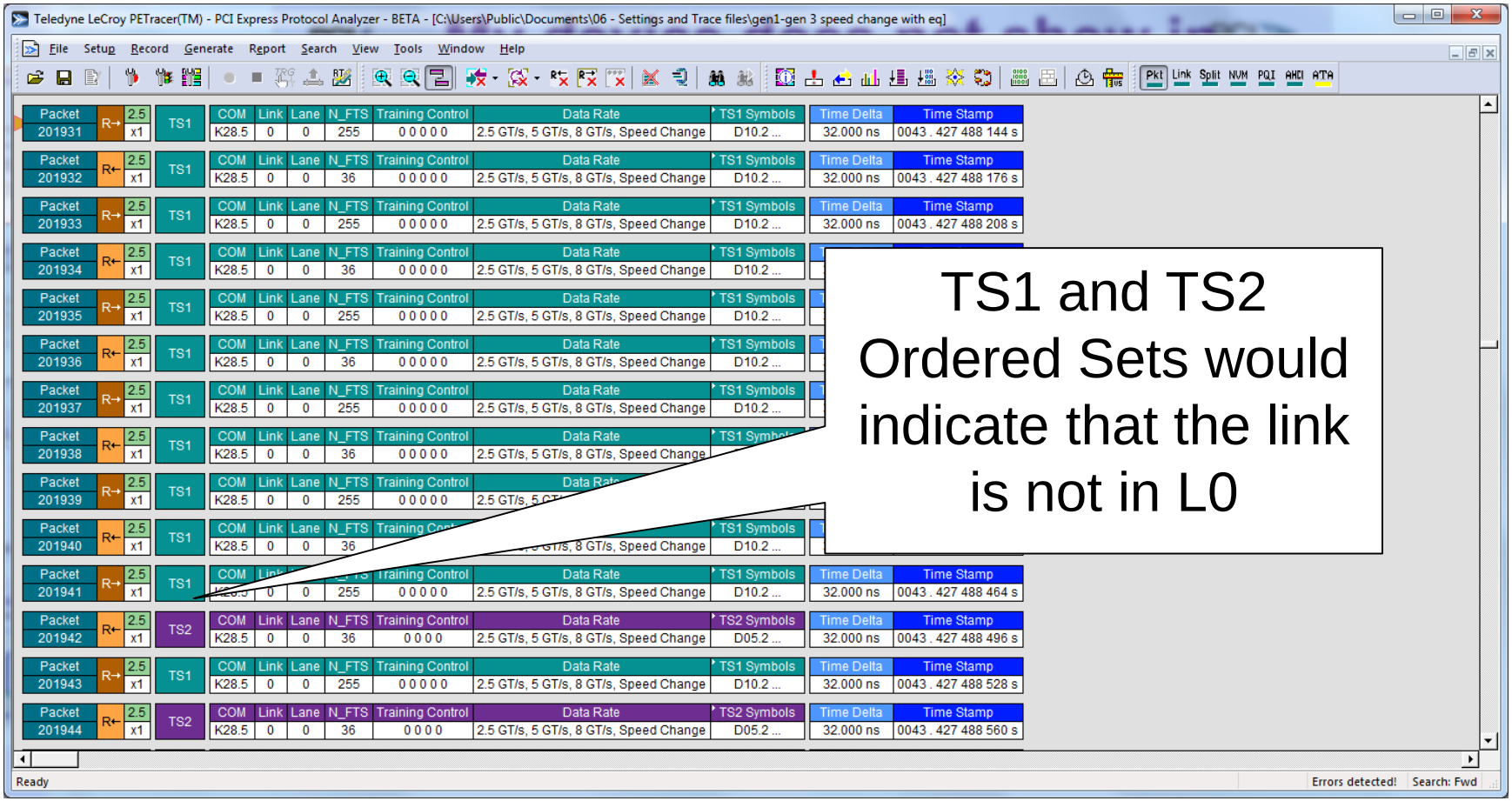
- **Is the Signal Integrity of the Link good on both sides**
- **Signal Integrity problems can show in a whole manner of ways, from instable links to devices just not showing up**





Copyright © 2017 PCI-SIG® - All Rights Reserved

Snapshot Trigger



Copyright © 2017 PCI-SIG® - All Rights Reserved

PCIe Link Training From the Specification

LTSSM State	Link Width	Link Speed	LinkUp	Link Training	Receiver Error	In-Band Presence[44]
Loopback	No action	Link speed may change on entry to Loopback from Configuration.	0b	0b	No action	1b
Hot Reset	No action	No action	0b	0b	Optional: Set on 8b/10b Error	1b

The state machine rules for configuring and operating a PCI Express Link are defined in the following sections.

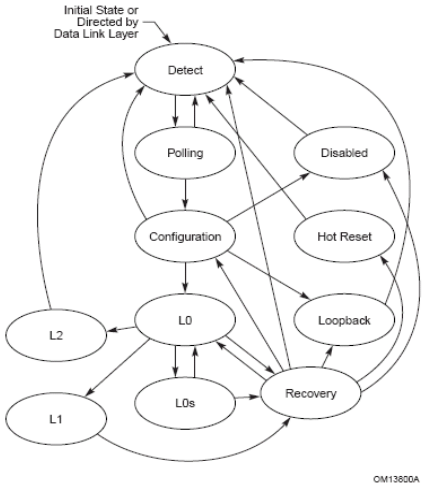


Figure 4-11: Main State Diagram for Link Training and Status State Machine

4.2.6.1. Detect

The Detect substate machine is shown in Figure 4-12.

4.2.6.1.1. Detect.Quiet

- ❑ Transmitter is in an Electrical Idle state.
  - Note: The DC common mode voltage is not required to be within specification.
- ❑ 2.5 GT/s data rate is selected as the frequency of operation. If the frequency of operation was not 2.5 GT/s data rate on entry to this substate, the LTSSM must stay in this substate for at least 1 ms, during which the frequency of operation must be changed to the 2.5 GT/s data rate.
  - Note: This does not affect the advertised data rate in the TS1 and TS2 Ordered Sets.
- ❑ LinkUp = 0b (status is cleared).
- ❑ The directed_speed_change variable is reset to 0b. The upconfigure_capable variable is reset to 0b. The idle_to_rlock_transitioned variable is reset to 0b. The select_deemphasis variable must be set to either 0b or 1b based on platform specific needs for the Downstream component and identical to the Selectable De-emphasis bit in the Link Control 2 register for an Upstream component.
  - Note that since these variables are defined with the 2.0 specification, pre-2.0 devices would not implement these variables and will always take the path as if the directed_speed_change and upconfigure_capable variables are constantly reset to 0b and the idle_to_rlock_transitioned variable is constantly set to 1b.
- ❑ The next state is Detect.Active after a 12 ms timeout or if Electrical Idle is broken on any Lane.

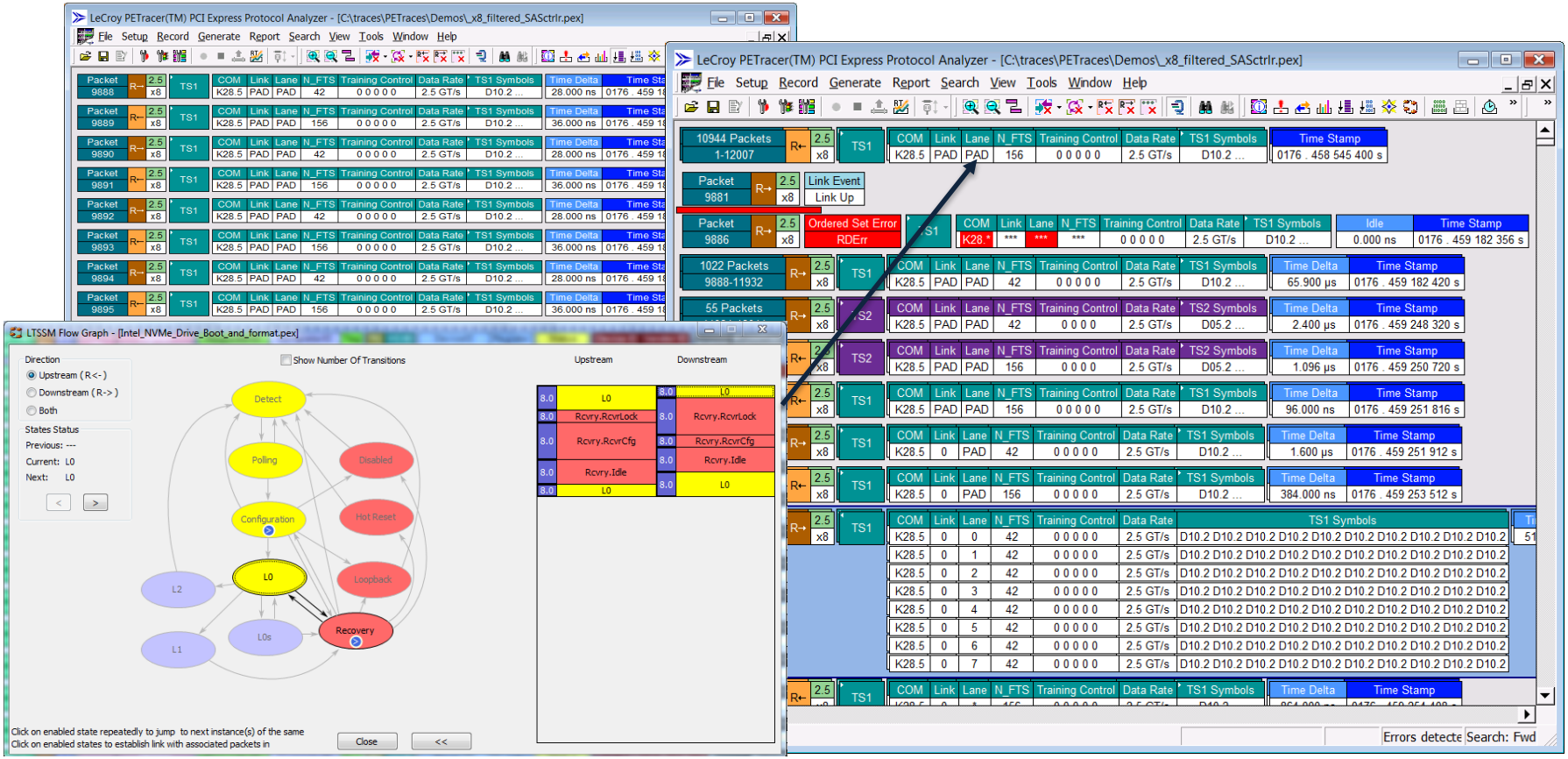
4.2.6.1.2. Detect.Active

- ❑ The Transmitter performs a Receiver Detection sequence on all un-configured Lanes that can form one or more Links (see Section 4.3.1.8 for more information).
- ❑ Next state is Polling if a Receiver is detected on all unconfigured Lanes.
- ❑ Next state is Detect.Quiet if a Receiver is not detected on any Lanes.

Copyright © 2017 PCI-SIG® - All Rights Reserved

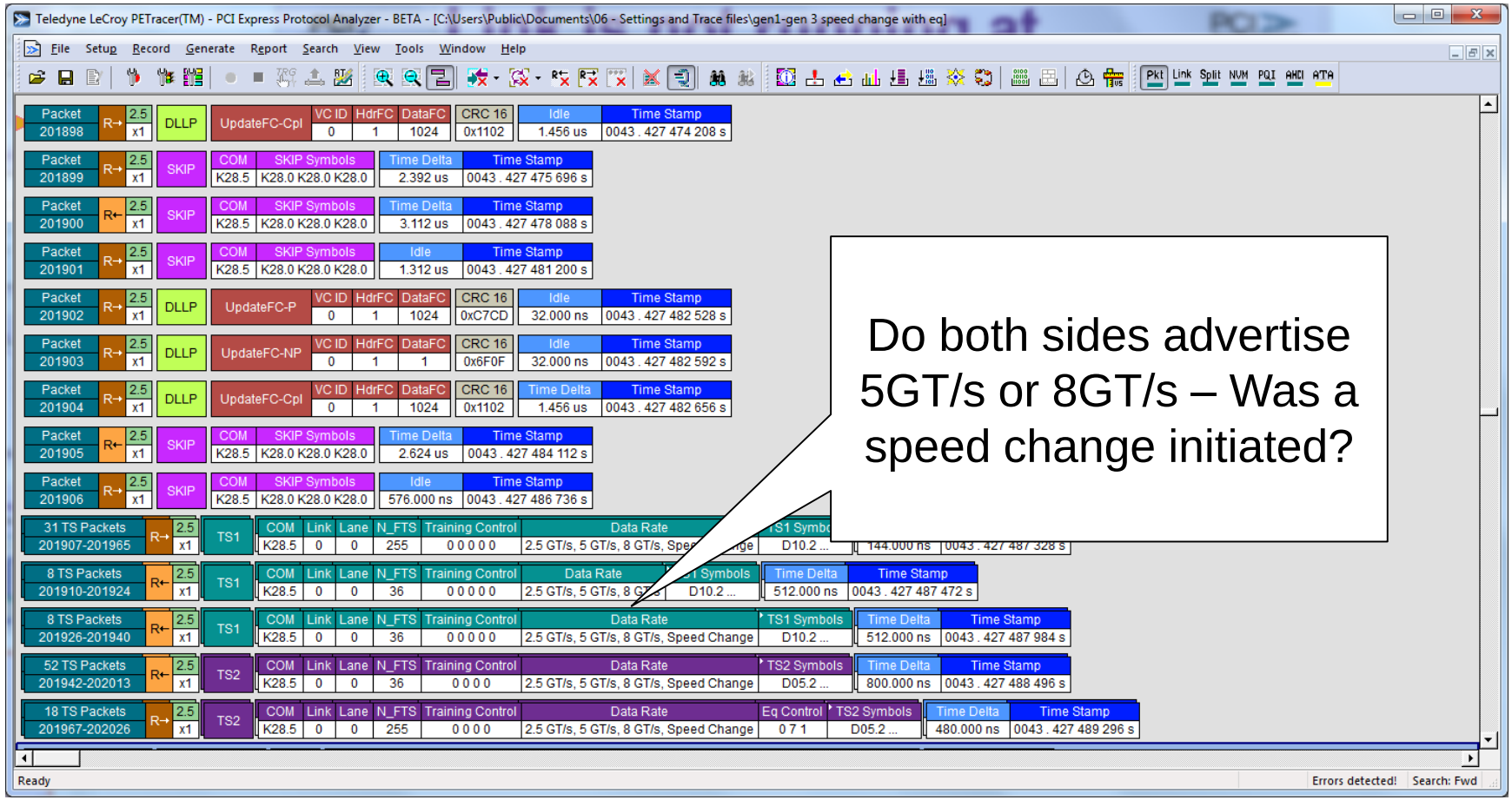
PCIe Link Training Debug

- **LTSSM state changes appear as 1000's of ordered sets**



Copyright © 2017 PCI-SIG® - All Rights Reserved

Example:
Link is Not Running at Maximum Speed



Copyright © 2017 PCI-SIG® - All Rights Reserved

Dynamic Equalization

- **The PCI Express 3.0 Specification allows for 8.0GT/s data rate**
- **In order to establish a link reliably at 8GT/s, the protocol allows for dynamic equalization of the link during the speed change in the Recovery state**
- **This is a 4 phase process and happens in the Recovery.Equalization substate**

Copyright © 2017 PCI-SIG® - All Rights Reserved

Recovery Substate Machine

Recovery.Equalization Substate

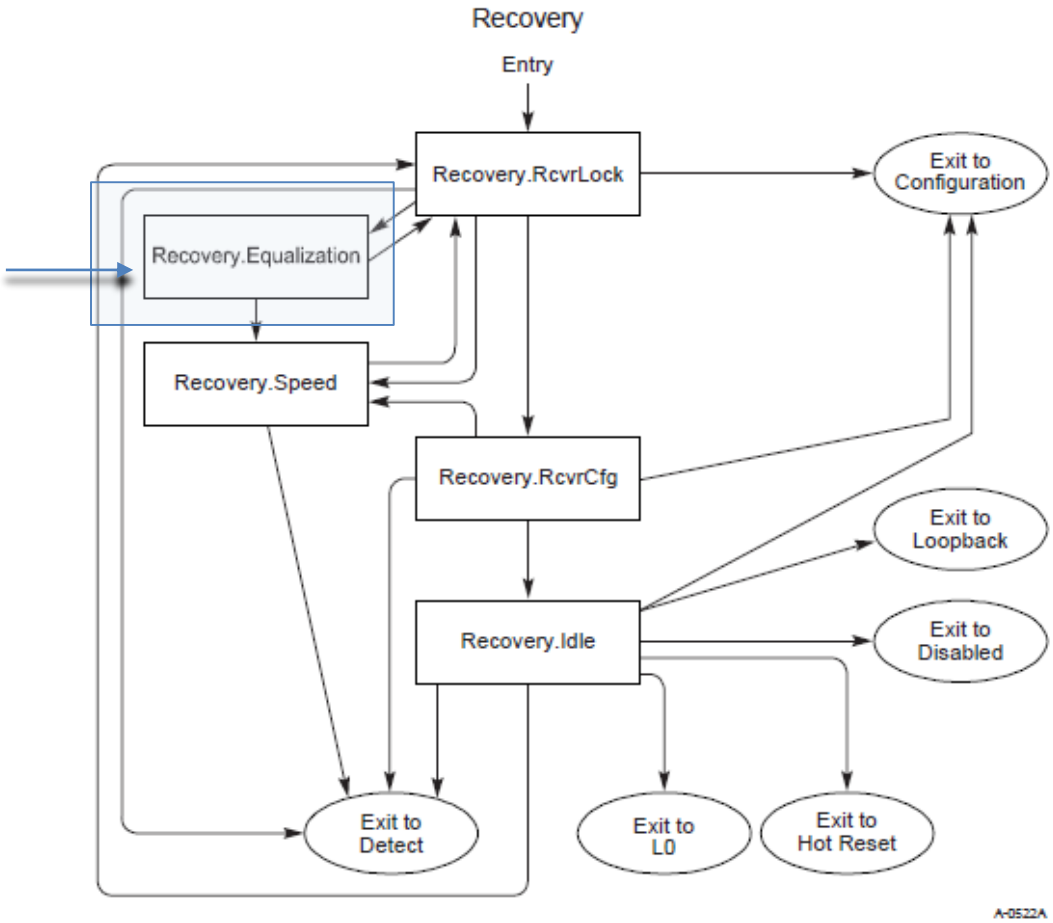


Figure 4-26: Recovery Substate Machine

Copyright © 2017 PCI-SIG® - All Rights Reserved

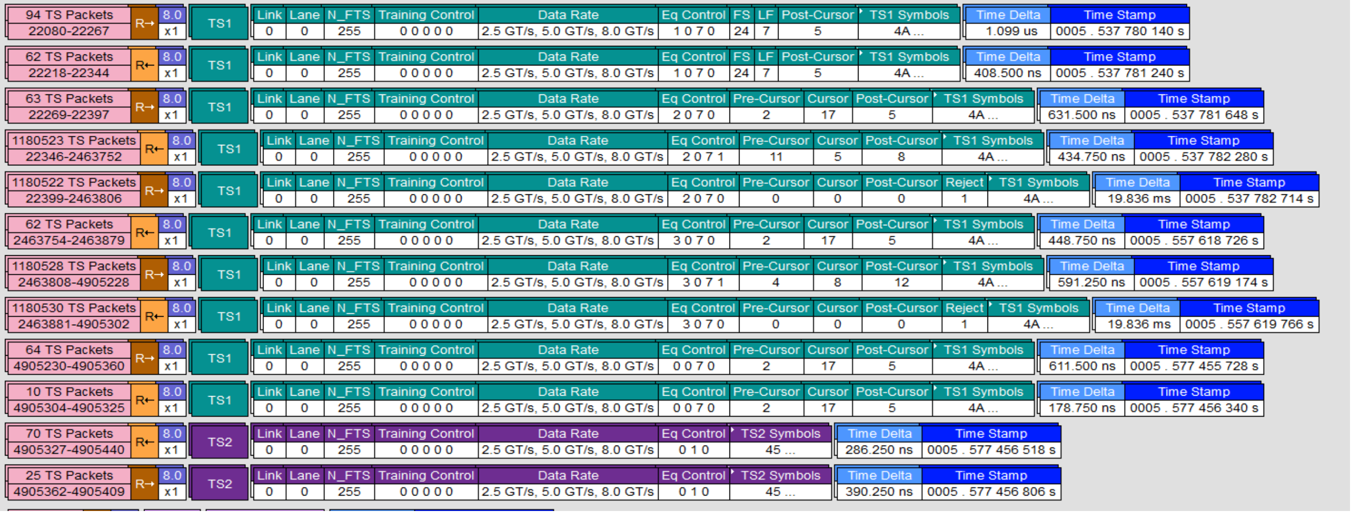
Equalization Process

- **See Section 4.2.3 of PCIe 4.0 Base Specification**
- **The Link equalization procedure enables components to:**
  - Adjust the Transmitter and the Receiver setup of each Lane to improve the signal quality
  - Meet the requirements specified in the physical layer specification, when operating at 8.0 GT/s and higher data rates.
- **The procedure must be executed during the first data rate change to 8.0 GT/s as well as the first change to all data rates greater than 8.0 GT/s.**
- **Components must not require that the equalization procedure be repeated at any data rate for reliable operation, although there is provision to repeat the procedure.**

Copyright © 2017 PCI-SIG® - All Rights Reserved

Equalization Process Part 1
2.5GT/s – 8GT/s

- **Process to change to 16GT/s starts with a speed change to 8GT/s, without advertising 16GT/S**
- **“The procedure must be executed during the first data rate change to 8.0 GT/s as well as the first change to all data rates greater than 8.0 GT/s.”**



				Link	Lane	N_FTS	Training Control	Data Rate	Eq Control	FS	LF	Pre-Cursor	Cursor	Post-Cursor	Reject	TS1/TS2 Symbols	Time Delta	Time Stamp
94 TS Packets 22080-22267	R→	8.0 x1	TS1	0	0	255	0 0 0 0 0	2.5 GT/s, 5.0 GT/s, 8.0 GT/s	1 0 7 0	24	7			5		4A ...	1.099 us	0005 . 537 780 140 s
62 TS Packets 22218-22344	R←	8.0 x1	TS1	0	0	255	0 0 0 0 0	2.5 GT/s, 5.0 GT/s, 8.0 GT/s	1 0 7 0	24	7			5		4A ...	408.500 ns	0005 . 537 781 240 s
63 TS Packets 22269-22397	R→	8.0 x1	TS1	0	0	255	0 0 0 0 0	2.5 GT/s, 5.0 GT/s, 8.0 GT/s	2 0 7 0			2	17	5		4A ...	631.500 ns	0005 . 537 781 648 s
1180523 TS Packets 22346-2463752	R←	8.0 x1	TS1	0	0	255	0 0 0 0 0	2.5 GT/s, 5.0 GT/s, 8.0 GT/s	2 0 7 1			11	5	8		4A ...	434.750 ns	0005 . 537 782 280 s
1180522 TS Packets 22399-2463806	R→	8.0 x1	TS1	0	0	255	0 0 0 0 0	2.5 GT/s, 5.0 GT/s, 8.0 GT/s	2 0 7 0			0	0	0	1	4A ...	19.836 ms	0005 . 537 782 714 s
62 TS Packets 2463754-2463879	R←	8.0 x1	TS1	0	0	255	0 0 0 0 0	2.5 GT/s, 5.0 GT/s, 8.0 GT/s	3 0 7 0			2	17	5		4A ...	448.750 ns	0005 . 557 618 726 s
1180528 TS Packets 2463808-4905228	R→	8.0 x1	TS1	0	0	255	0 0 0 0 0	2.5 GT/s, 5.0 GT/s, 8.0 GT/s	3 0 7 1			4	8	12		4A ...	591.250 ns	0005 . 557 619 174 s
1180530 TS Packets 2463881-4905302	R←	8.0 x1	TS1	0	0	255	0 0 0 0 0	2.5 GT/s, 5.0 GT/s, 8.0 GT/s	3 0 7 0			0	0	0	1	4A ...	19.836 ms	0005 . 557 619 766 s
64 TS Packets 4905230-4905360	R→	8.0 x1	TS1	0	0	255	0 0 0 0 0	2.5 GT/s, 5.0 GT/s, 8.0 GT/s	0 0 7 0			2	17	5		4A ...	611.500 ns	0005 . 577 455 728 s
10 TS Packets 4905304-4905325	R←	8.0 x1	TS1	0	0	255	0 0 0 0 0	2.5 GT/s, 5.0 GT/s, 8.0 GT/s	0 0 7 0			2	17	5		4A ...	178.750 ns	0005 . 577 456 340 s
70 TS Packets 4905327-4905440	R←	8.0 x1	TS2	0	0	255	0 0 0 0 0	2.5 GT/s, 5.0 GT/s, 8.0 GT/s	0 1 0							45 ...	286.250 ns	0005 . 577 456 518 s
25 TS Packets 4905362-4905409	R→	8.0 x1	TS2	0	0	255	0 0 0 0 0	2.5 GT/s, 5.0 GT/s, 8.0 GT/s	0 1 0							45 ...	390.250 ns	0005 . 577 456 806 s

Copyright © 2017 PCI-SIG® - All Rights Reserved

Equalization Process Part 2
8GT/s – 16GT/s

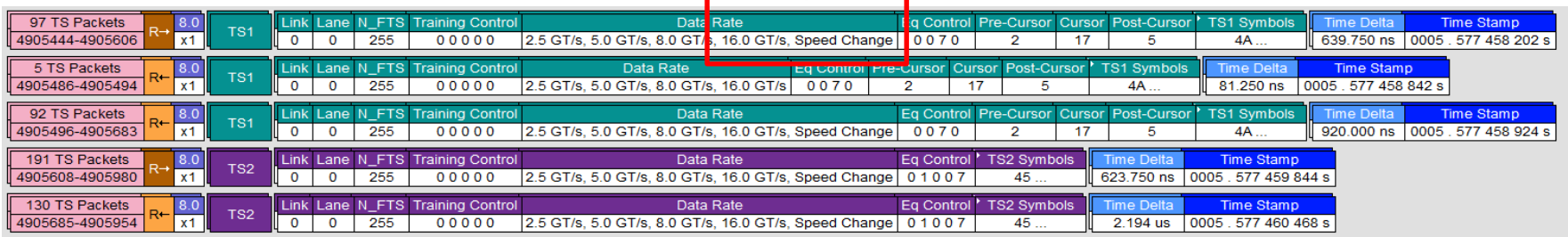
- **Immediately following the transition from Recovery to L0, after the initial data rate change to 8.0 GT/s, the Downstream Port is required to:**
  - transition from L0 to Recovery
  - advertise 16.0 GT/s data rate support
  - change the data rate to 16.0 GT/s and perform the 16.0 GT/s equalization procedure
- **If**
  - both components advertised that they are capable of 16.0 GT/s during the initial Link negotiation,
  - neither component detected problems with its 8.0 GT/s equalization settings and it intends to perform a 16.0 GT/s equalization procedure using the autonomous mechanism.

Copyright © 2017 PCI-SIG® - All Rights Reserved

Equalization Process Part 2
8GT/s – 16GT/s

- **Speed Change is initiated to change to speed above 8GT/s**

TS1 advertises 16.0GT/s Data Rate



97 TS Packets 4905444-4905606	R→	8.0 x1	TS1	Link	Lane	N_FTS	Training Control	Data Rate	Eq Control	Pre-Cursor	Cursor	Post-Cursor	TS1 Symbols	Time Delta	Time Stamp				
				0	0	255	0 0 0 0 0	2.5 GT/s, 5.0 GT/s, 8.0 GT/s, 16.0 GT/s, Speed Change	0 0 7 0	2	17	5	4A ...	639.750 ns	0005 . 577 458 202 s				
5 TS Packets 4905486-4905494	R←	8.0 x1	TS1	Link	Lane	N_FTS	Training Control	Data Rate	Eq Control	Pre-Cursor	Cursor	Post-Cursor	TS1 Symbols	Time Delta	Time Stamp				
				0	0	255	0 0 0 0 0	2.5 GT/s, 5.0 GT/s, 8.0 GT/s, 16.0 GT/s	0 0 7 0	2	17	5	4A ...	81.250 ns	0005 . 577 458 842 s				
92 TS Packets 4905496-4905683	R←	8.0 x1	TS1	Link	Lane	N_FTS	Training Control	Data Rate	Eq Control	Pre-Cursor	Cursor	Post-Cursor	TS1 Symbols	Time Delta	Time Stamp				
				0	0	255	0 0 0 0 0	2.5 GT/s, 5.0 GT/s, 8.0 GT/s, 16.0 GT/s, Speed Change	0 0 7 0	2	17	5	4A ...	920.000 ns	0005 . 577 458 924 s				
191 TS Packets 4905608-4905980	R→	8.0 x1	TS2	Link	Lane	N_FTS	Training Control	Data Rate	Eq Control	TS2 Symbols	Time Delta	Time Stamp							
				0	0	255	0 0 0 0 0	2.5 GT/s, 5.0 GT/s, 8.0 GT/s, 16.0 GT/s, Speed Change	0 1 0 0 7	45 ...	623.750 ns	0005 . 577 459 844 s							
130 TS Packets 4905685-4905954	R←	8.0 x1	TS2	Link	Lane	N_FTS	Training Control	Data Rate	Eq Control	TS2 Symbols	Time Delta	Time Stamp							
				0	0	255	0 0 0 0 0	2.5 GT/s, 5.0 GT/s, 8.0 GT/s, 16.0 GT/s, Speed Change	0 1 0 0 7	45 ...	2.194 us	0005 . 577 460 468 s							

Copyright © 2017 PCI-SIG® - All Rights Reserved

Potential Equalization Problems

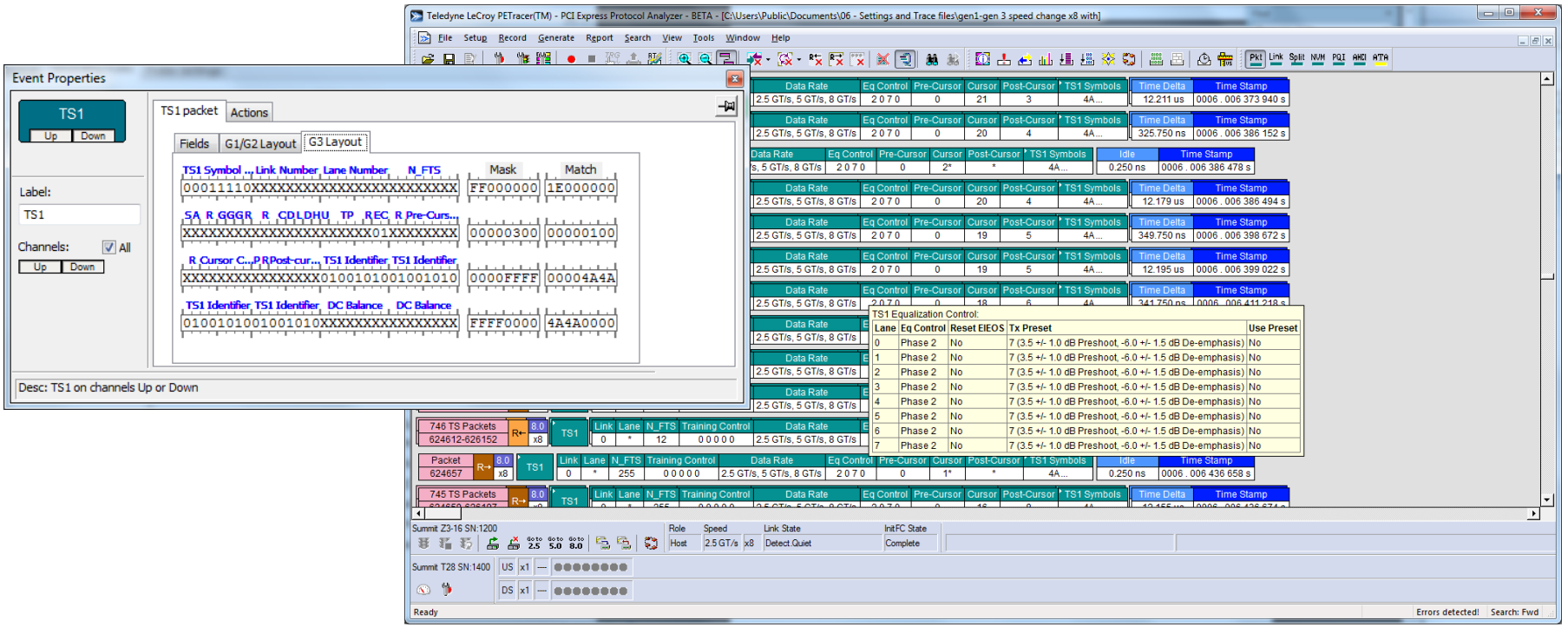
- **Link does not get to 8GT/s or 16GT/s**
- **Link gets to 8GT/s or 16GT/s but is unstable**
- **Equalization fails because of wrong coefficients**
- **Does the equalization process complete?**
- **Protocol Analyzer is the only way to get visibility**



Copyright © 2017 PCI-SIG® - All Rights Reserved

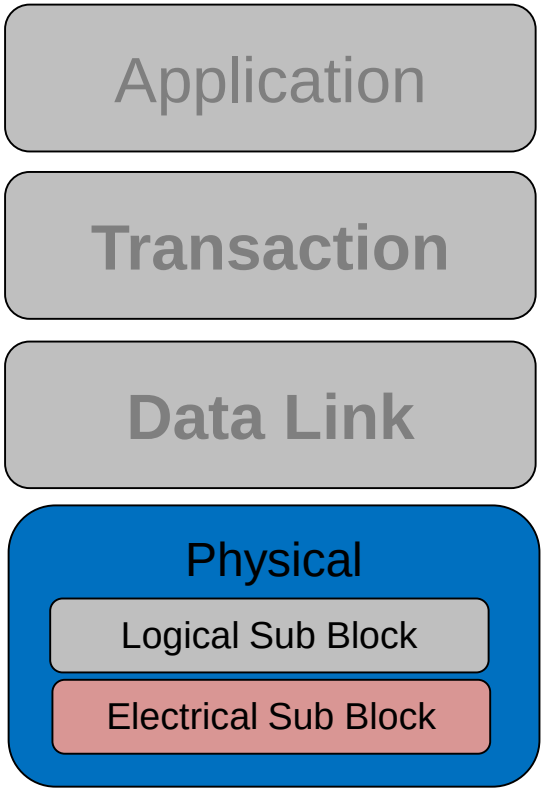
How Do I Debug Equalization?

- **Use the EC field as a trigger – this will help navigate the equalization**
- **Compress the Training Sequences helps digest the process – there are thousands of them!**



Copyright © 2017 PCI-SIG® - All Rights Reserved

Polarity Issues



Copyright © 2017 PCI-SIG® - All Rights Reserved

Polarity Inversion

- **PCI Express allows polarity inversion**
- **New form factors and adapters are exhibiting this more often**
- **This can often be challenging to work out when setting up an analyzer.**
- **It is quite common to see different polarity settings on different slots in the same system**

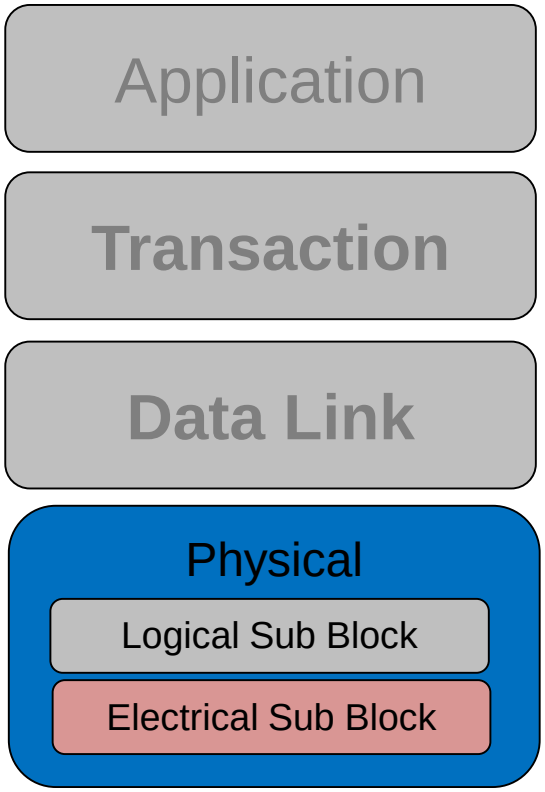
Copyright © 2017 PCI-SIG® - All Rights Reserved

Polarity Issues

- **Depending on the type of probe being used, this may change the Polarity settings on the link**
- **Normally Polarity can be auto detected**
  - It is fixed for any particular slot or card
  - If Power management is enabled, it is always better to use manually assigned polarity settings
- **Incorrect Polarity settings will result in a trace full of errors**

Copyright © 2017 PCI-SIG® - All Rights Reserved

Lane Reversal



Copyright © 2017 PCI-SIG® - All Rights Reserved

Lane Reversal

- **Many implementations are now using Lane Reversal**
- **Especially true in Server platforms with U.2 connections**
- **This often causes analysis problems**
- **How to identify Lane Reversal?**

Copyright © 2017 PCI-SIG® - All Rights Reserved

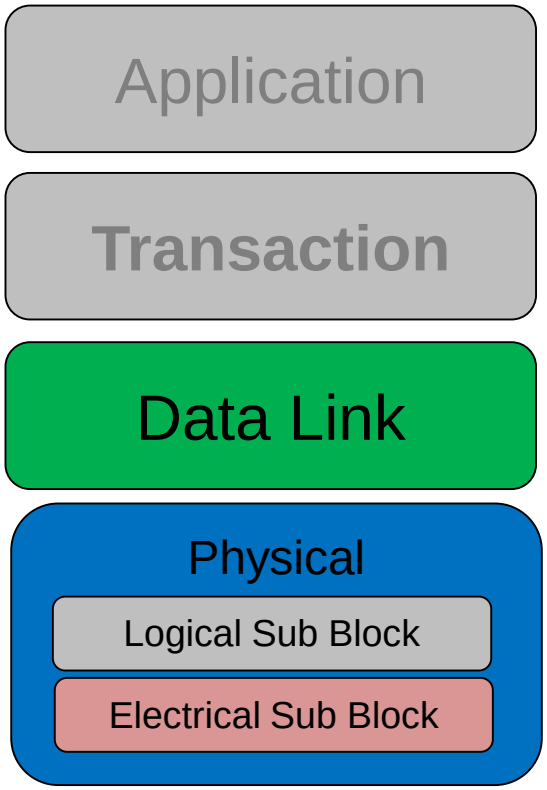
Lane Reversal

- **The trick when looking for Lane Reversal is:**
  - Trigger on TS2 Ordered Set
  - Look in the Lane Number field
  - This should read 0,1,2,3
  - If it reads 3,2,1,0 then the link has Lane Reversal

Packet				COM	Lin	Lane	N_FTS	Training Control	Data Rate	TS2 Symbols	Time Delta	Time Stamp
4437	R←	2.5 x4	TS2	K28.5	0	0	255	0 0 0 0	2.5 GT/s, 5 GT/s, 8 GT/s	D05.2 D05.2 D05.2 D05.2 D05.2 D05.2 D05.2 D05.2 D05.2 D05.2	0.000 ns	0007 . 165 541 810 s
				K28.5	0	1	255	0 0 0 0	2.5 GT/s, 5 GT/s, 8 GT/s	D05.2 D05.2 D05.2 D05.2 D05.2 D05.2 D05.2 D05.2 D05.2 D05.2		
				K28.5	0	2	255	0 0 0 0	2.5 GT/s, 5 GT/s, 8 GT/s	D05.2 D05.2 D05.2 D05.2 D05.2 D05.2 D05.2 D05.2 D05.2 D05.2		
				K28.5	0	3	255	0 0 0 0	2.5 GT/s, 5 GT/s, 8 GT/s	D05.2 D05.2 D05.2 D05.2 D05.2 D05.2 D05.2 D05.2 D05.2 D05.2		

Copyright © 2017 PCI-SIG® - All Rights Reserved

Power Management Problems



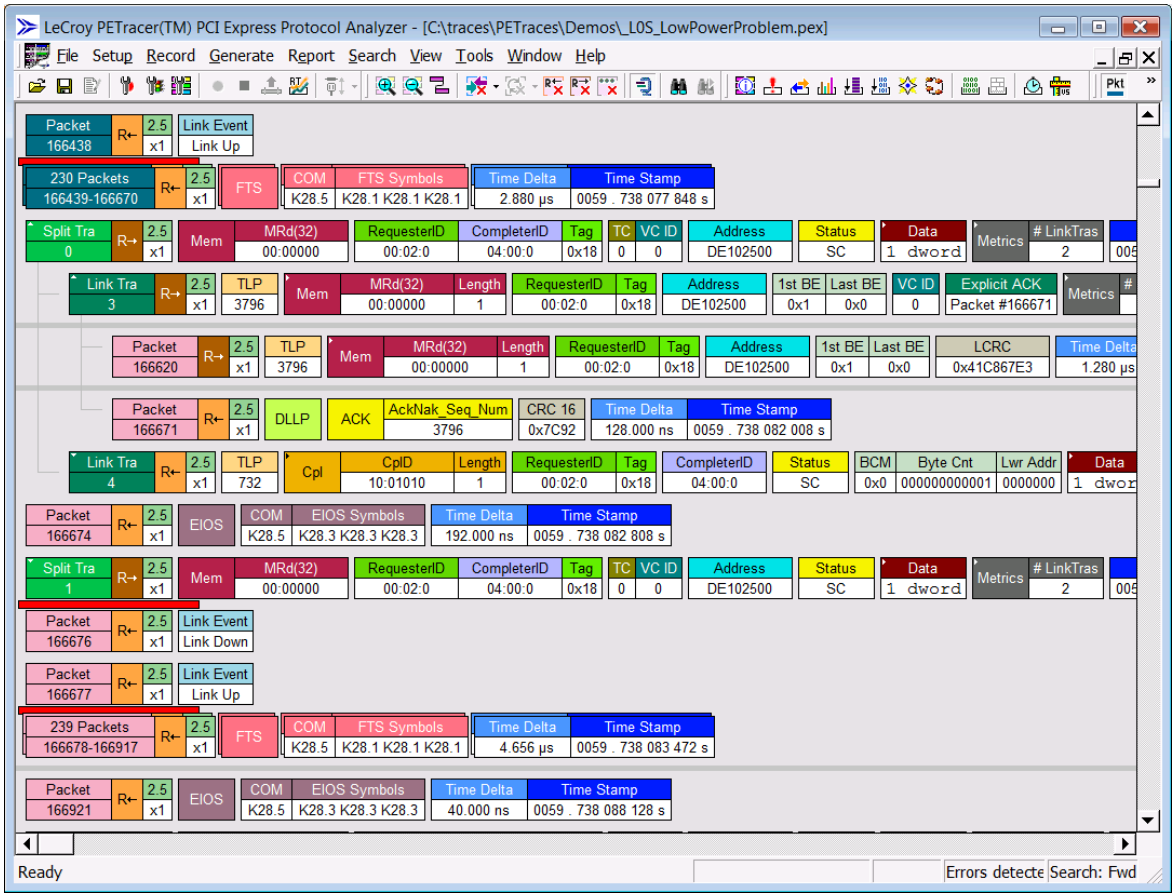
Copyright © 2017 PCI-SIG® - All Rights Reserved

L0s

- **It can be challenging to debug L0s problems, it is important to set things up properly**
  - The exit latency from L0s is very short
  - Give the analyzer the best chance of locking by disabling auto detection such as speed, link width and polarity
  - Make sure the analyzer is set up properly with L0s turned off and ensure clean captures

Copyright © 2017 PCI-SIG® - All Rights Reserved

ASPM – L0s



Copyright © 2017 PCI-SIG® - All Rights Reserved

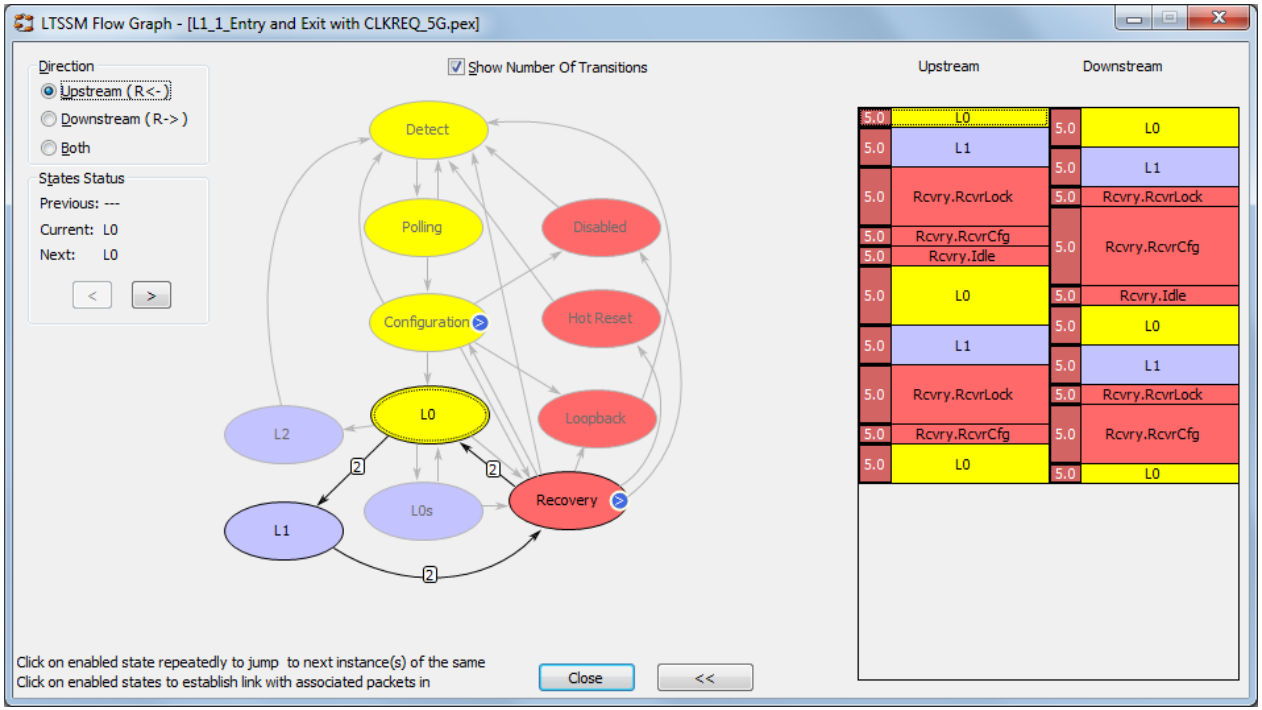
L1 and L1 Substates

- **L1 is a more aggressive form of power management than L0s, and the L1 Substates ECN takes that further**
- **There are 2 mechanisms to enter L1 - ASPM L1 and PCIPM L1**
  - Each have a different entry mechanism
  - PCIPM uses a config write from the RC to the PMSCR register on the device to initiate the transition to L1
  - ASPM L1 the downstream component indicates the desire to enter the L1 state, and sends PM_Active_State_Request_L1 to the root, the root acknowledges with PM_Request_ACK DLLP

Copyright © 2017 PCI-SIG® - All Rights Reserved

ASPM – L1

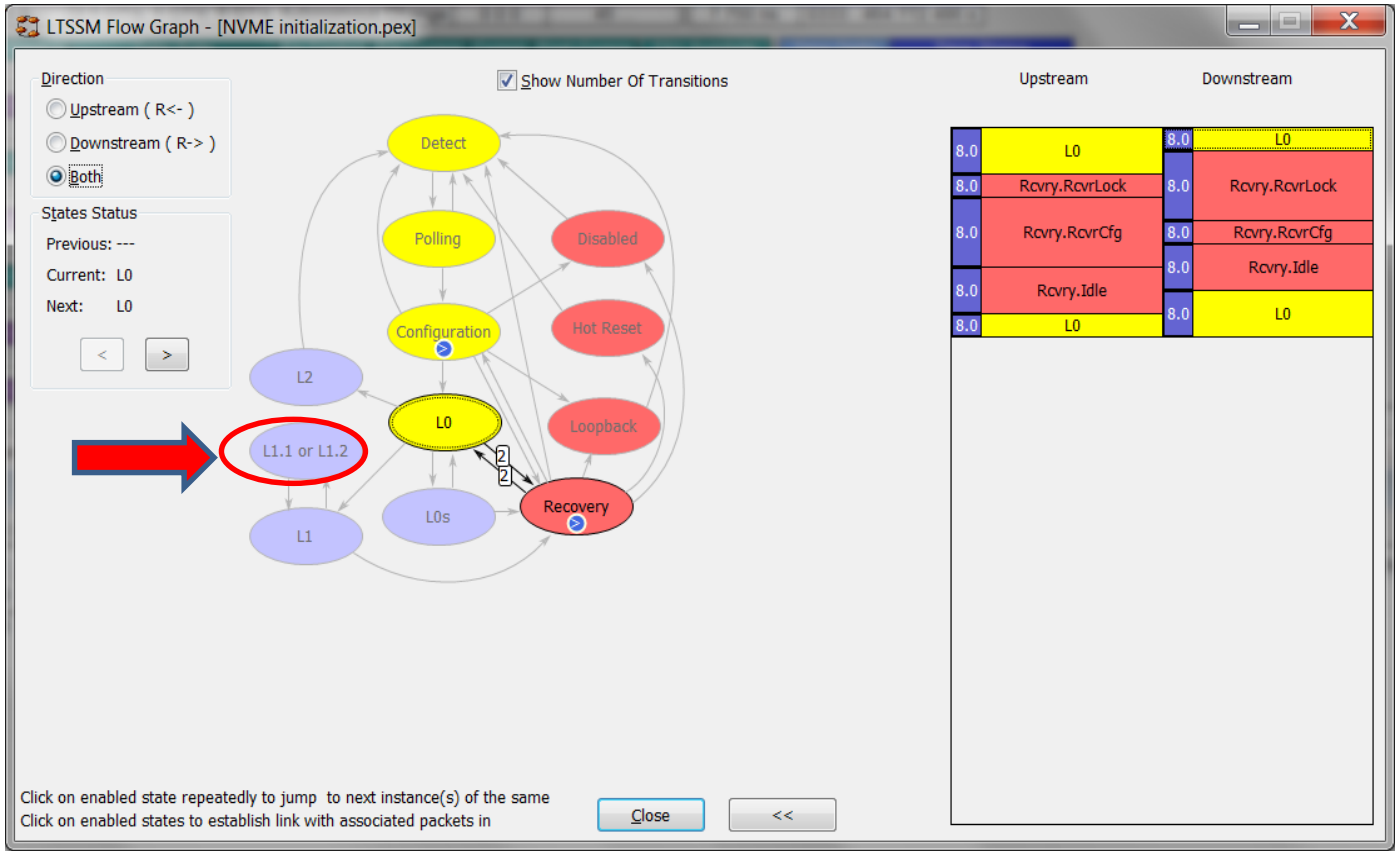
- **Advanced post processing shows how many times each state was entered**



Copyright © 2017 PCI-SIG® - All Rights Reserved

ASPM – L1.1 and L1.2

- **L1.1 and L1.2 are substates of L1**



Copyright © 2017 PCI-SIG® - All Rights Reserved

L1 Substates

- **A device indicates its support of L1 substates in the configuration space**
- **L1 PM Substates is considered enabled on a Port when any combination of the ASPM L1.1 Enable, ASPM L1.2 Enable, PCI-PM L1.1 Enable and PCI-PM L1.2 Enable bits associated with that Port are Set**

Copyright © 2017 PCI-SIG® - All Rights Reserved

How Does it Look on a Protocol Analyzer?

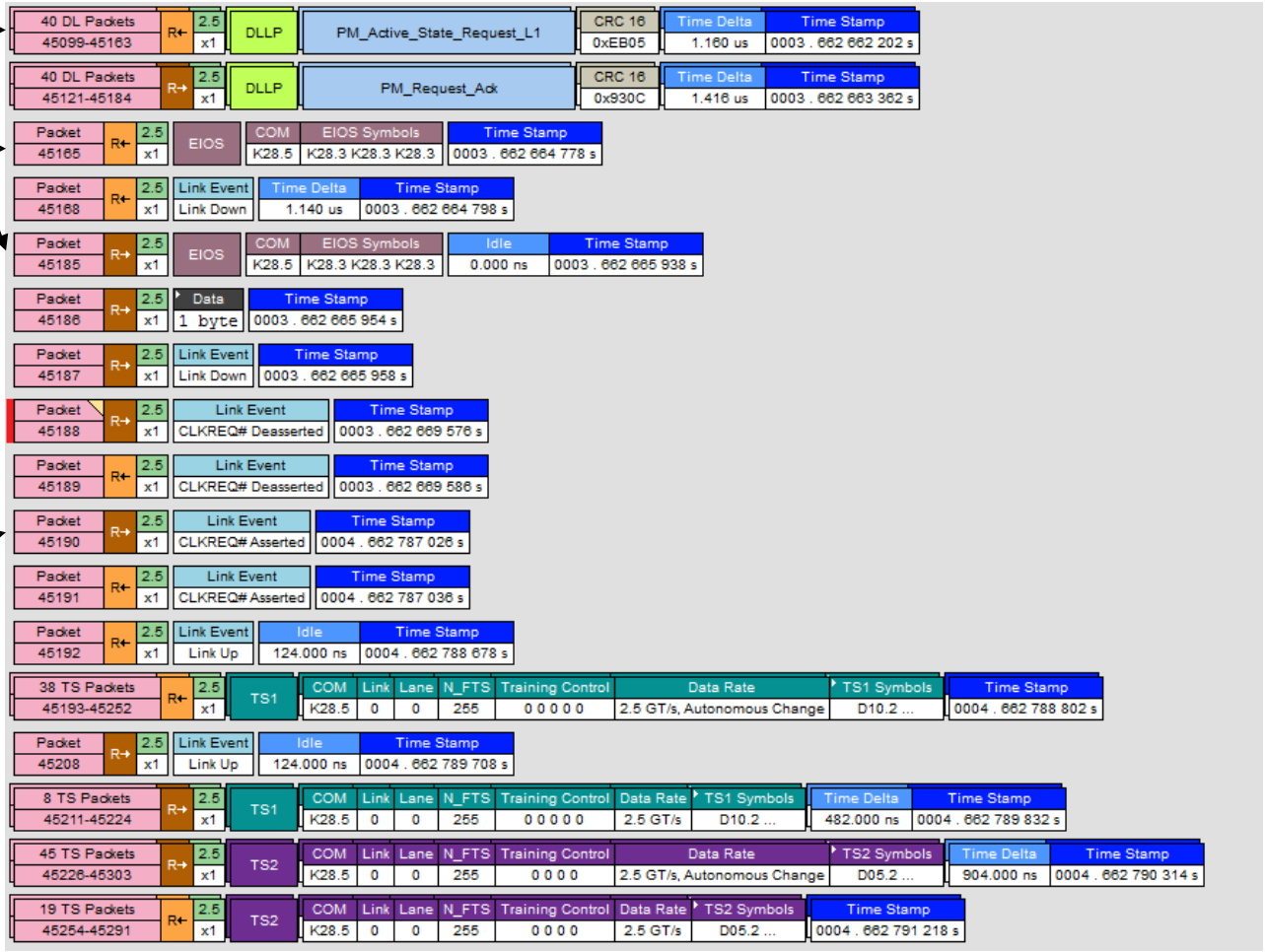
L1 Entry Request

Link Enters Electrical Idle

CLKREQ# Deasserted

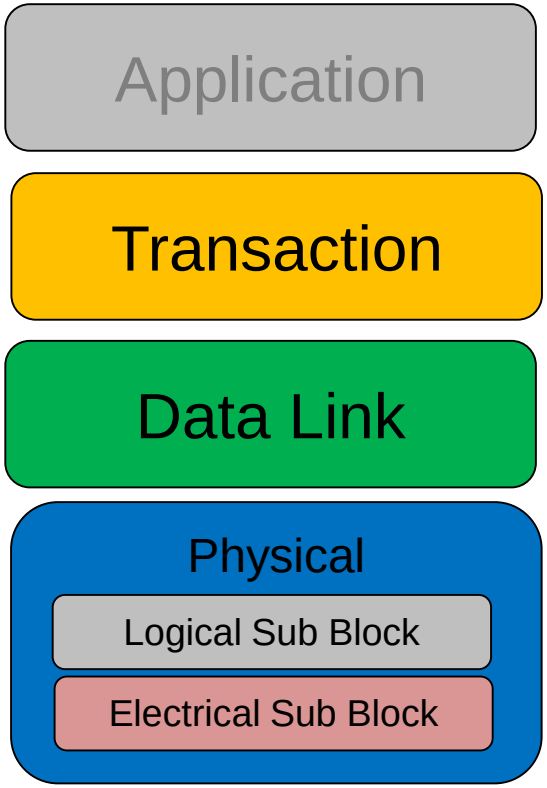
CLKREQ# Asserted

Link Enters Recovery



Copyright © 2017 PCI-SIG® - All Rights Reserved

Reset Problems



Copyright © 2017 PCI-SIG® - All Rights Reserved

PCIe Reset Mechanisms

- **There are different types of reset mechanism**
  - Cold, Warm, Hot reset
  - Function Level reset

Copyright © 2017 PCI-SIG® - All Rights Reserved

Reset Mechanisms

- **Conventional Reset includes all reset mechanisms excluding Function Level Reset**
- **All form factors need a mechanism to return all Port state to the initial condition, called Fundamental Reset**
  - Must occur following application of power to the component – called a “Cold reset”
  - Can be done without removing power – “Warm reset”
  - In band mechanism for propagating Conventional reset across a Link – “Hot reset”
- **See section 6.6.1 in PCI Express Specification**

Copyright © 2017 PCI-SIG® - All Rights Reserved

PCIe Reset Mechanisms - FLR

- **Function Level Reset**
- **The FLR mechanism enables software to quiesce and reset Endpoint hardware with Function-level granularity**
- **FLR applies on a per Function basis. Only the targeted Function is affected by the FLR operation.**
- **The Link state must not be affected by an FLR.**
- **Note that the controls that enable the Function to initiate requests on PCI Express are cleared, including Bus Master Enable, MSI Enable, and the like, effectively causing the Function to become quiescent on the Link.**
- **After an FLR has been initiated by writing a 1b to the Initiate Function Level Reset bit, the Function 15 must complete the FLR within 100 ms.**
- **FLR is optional in PCIe spec but required in NVMe**

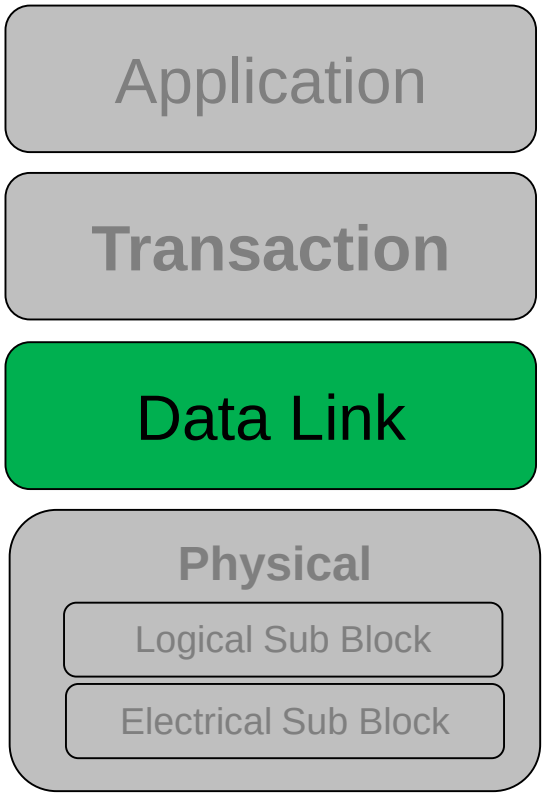
Copyright © 2017 PCI-SIG® - All Rights Reserved

Typical Problems with Reset

- **Possible problems that may occur**
  - Registers not reset properly
  - Errors reported during Reset
  - Link fails after Reset

Copyright © 2017 PCI-SIG® - All Rights Reserved

Flow Control Problems



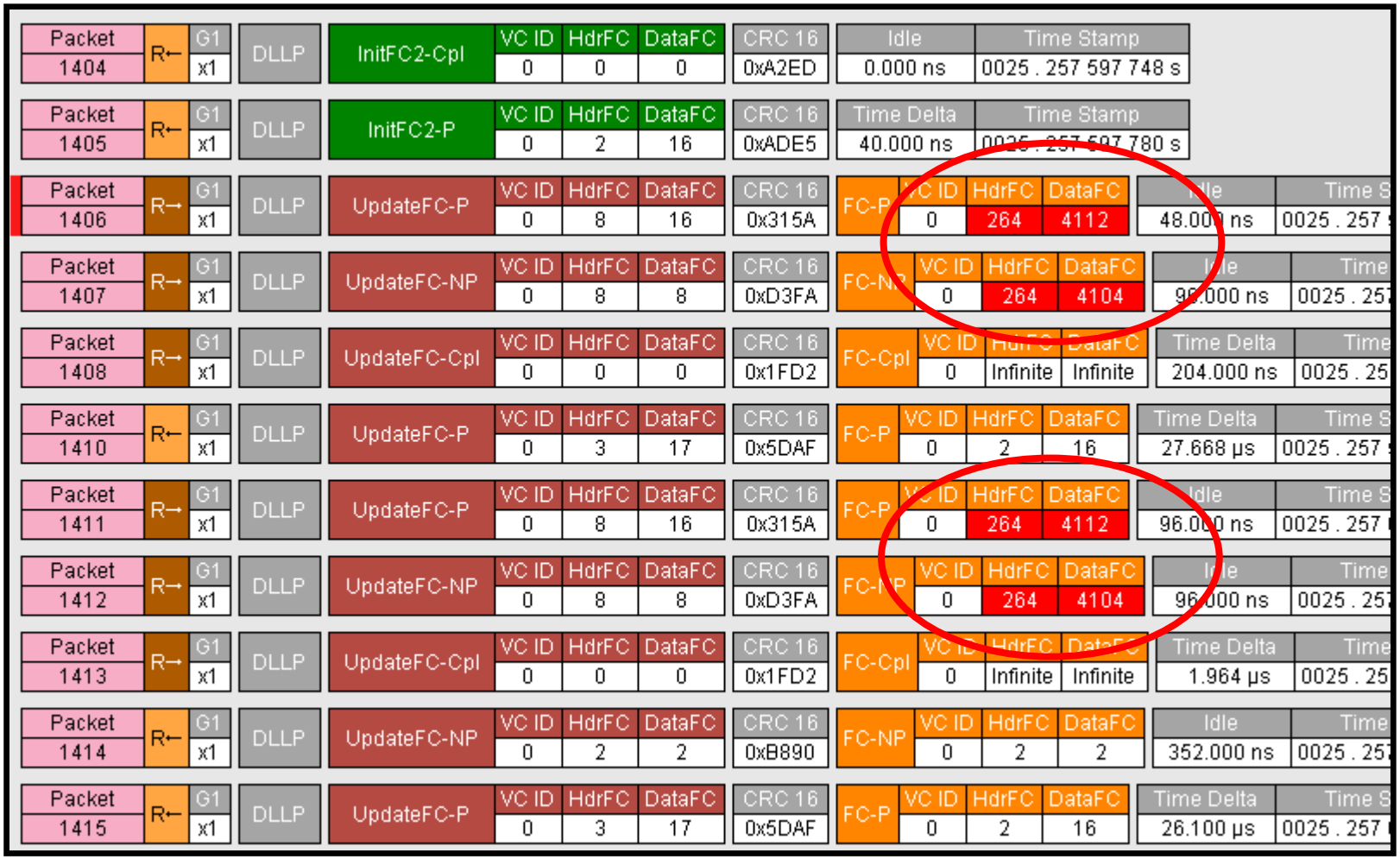
Copyright © 2017 PCI-SIG® - All Rights Reserved

Data Link Layer Issues

- **Some problems are related to flow control mechanisms**
  - Most DLL FC problems will not be discovered by the current compliance tests
  - Many DLL FC problems are not discovered until after a product is released
  - Problems may only show up with specific combination of Root and End Point that is not tested during compliance testing
  - Data Link Layer problems often result in performance issues

Copyright © 2017 PCI-SIG® - All Rights Reserved

Detecting Credit Errors



Packet		G1	DLLP		VC ID	HdrFC	DataFC	CRC 16					
1404	R←	x1	DLLP	InitFC2-Cpl	0	0	0	0xA2ED	Idle	0.000 ns	Time Stamp	0025 . 257 597 748 s	
1405	R←	x1	DLLP	InitFC2-P	0	2	16	0xADE5	Time Delta	40.000 ns	Time Stamp	0025 . 257 597 780 s	
1406	R←	x1	DLLP	UpdateFC-P	0	8	16	0x315A	FC-P	0	264	4112	48.000 ns	0025 . 257
1407	R←	x1	DLLP	UpdateFC-NP	0	8	8	0xD3FA	FC-NP	0	264	4104	96.000 ns	0025 . 257
1408	R←	x1	DLLP	UpdateFC-Cpl	0	0	0	0x1FD2	FC-Cpl	0	Infinite	Infinite	204.000 ns	0025 . 25
1410	R←	x1	DLLP	UpdateFC-P	0	3	17	0x5DAF	FC-P	0	2	16	27.668 µs	0025 . 257
1411	R←	x1	DLLP	UpdateFC-P	0	8	16	0x315A	FC-P	0	264	4112	96.000 ns	0025 . 257
1412	R←	x1	DLLP	UpdateFC-NP	0	8	8	0xD3FA	FC-NP	0	264	4104	96.000 ns	0025 . 257
1413	R←	x1	DLLP	UpdateFC-Cpl	0	0	0	0x1FD2	FC-Cpl	0	Infinite	Infinite	1.964 µs	0025 . 25
1414	R←	x1	DLLP	UpdateFC-NP	0	2	2	0xB890	FC-NP	0	2	2	352.000 ns	0025 . 257
1415	R←	x1	DLLP	UpdateFC-P	0	3	17	0x5DAF	FC-P	0	2	16	26.100 µs	0025 . 257

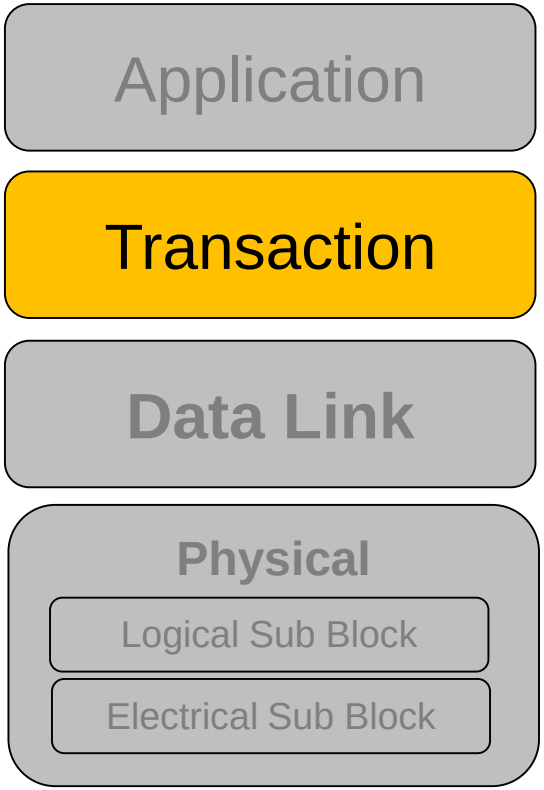
Copyright © 2017 PCI-SIG® - All Rights Reserved

Data Link Layer Problems

- **Flow control overflow**
- **Why is a packet getting a NAK?**
  - Set up trigger on NAK and look what happened before

Copyright © 2017 PCI-SIG® - All Rights Reserved

Class Code Issues

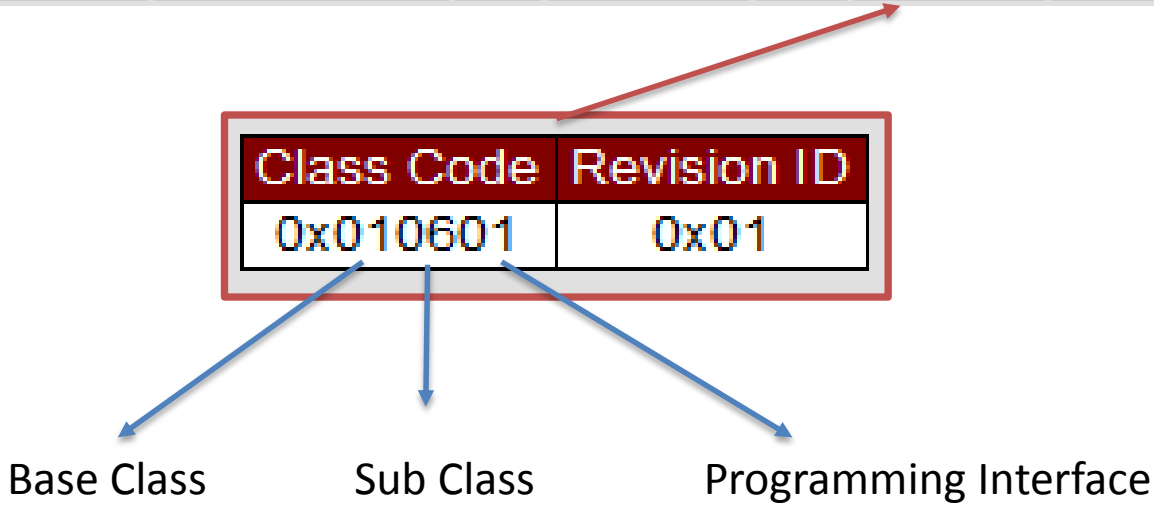


Copyright © 2017 PCI-SIG® - All Rights Reserved

Class Code

- **Class Code used to indicate type of device**

Split Tra	R→	5.0	Cfg	CfgRd0	RequesterID	CompleterID	Tag	TC	VC ID	DeviceID	Register	Status	Class Code	Revision ID	Metrics	# LinkTrans	Time Delta	Time Stamp
110		x1		000:00100	000:00:0	002:00:0	0	0	0	002:00:0	0x008	SC	0x010601	0x01		2	1.256 us	0006 . 122 324 510 s



Copyright © 2017 PCI-SIG® - All Rights Reserved

Class Code

Base Class	Sub-Class	Programming Interface	Meaning
01h	05h	20h	ATA controller with ADMA interface - single stepping (see Note 2)
		30h	ATA controller with ADMA interface - continuous operation (see Note 2)
	06h	00h	Serial ATA controller - vendor-specific interface
		01h	Serial ATA controller - AHCI interface (see note 7)
		02h	Serial Storage Bus Interface
	07h	00h	Serial Attached SCSI (SAS) controller - vendor-specific interface
		01h	Obsolete
	08h	00h	Non-volatile memory subsystem - vendor-specific interface
		01h	Non-volatile memory subsystem - NVMHCI interface (see note 8)
		02h	Non-volatile memory subsystem - NVM Express interface (see Note 6)
	09h	00h	Universal Flash Storage (UFS) controller - vendor-specific interface
		01h	Universal Flash Storage (UFS) controller - Universal Flash Storage Host Controller Interface (UFSHCI) (see Note 5)
	80h	00h	Other mass storage controller - vendor-specific interface

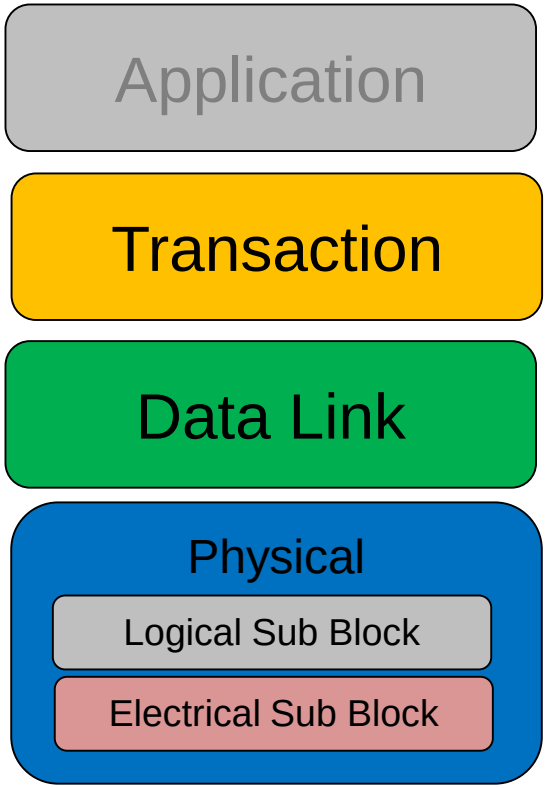
Copyright © 2017 PCI-SIG® - All Rights Reserved

Class Code Issues

- **Incorrect Class Code assignments cause problems**
- **Showing the device as the wrong type of device can cause the system to be unstable or crash**
- **Examples**
  - BIOS does not support it
  - Incorrect driver is used

Copyright © 2017 PCI-SIG® - All Rights Reserved

Compliance Test Failures



Copyright © 2017 PCI-SIG® - All Rights Reserved

“PCIe 1.0, 2.0, 3.0, 4.0???”

- **What is the difference?**
  - These are different revisions of the specification
  - They do not necessarily mean a device can support higher speeds
  - The changes go beyond the speed capability and include ECNs and other updates
- **We tend to use PCIe 1.x, PCIe 2.0, PCIe 3.0 synonymously with 2.5GT/s, 5GT/s and 8GT/s**
  - A PCIe 1.0a or 1.1 device supports 2.5GT/s
  - A PCIe 2.0 device must support 2.5GT/s and can support 5GT/s
  - A PCIe 3.0 device must support 2.5GT/s and can support up to 5GT/s and 8GT/s
  - A PCIe 4.0 device must support 2.5GT/s and can support up to 5GT/s, 8GT/s and 16GT/s
- **How does this apply to the layered model?**

Copyright © 2017 PCI-SIG® - All Rights Reserved

“PCIe 1.0, 2.0, 3.0, 4.0???”

- **The most significant change in specification revisions is to add higher speed capabilities**
- **However, there are many other ECNs that were approved after the release of a specification that are rolled into newer releases**
- **For example, there is a 3.1 revision of the specification that included ECNs proposed after the release of the 3.0 specification**
- **The PCI Express 4.0 Base Specification is currently being worked on**

Copyright © 2017 PCI-SIG® - All Rights Reserved

“PCIe 1.0, 2.0, 3.0, 4.0???”

- **There are many changes in Configuration Space that can results in differing test results**
- **For Example: The way PCI-SIG tests devices is not the same for a PCIe 1.1 device and a PCIe 3.0 device that only supports 2.5GT/s**

Copyright © 2017 PCI-SIG® - All Rights Reserved

PCI-SIG Compliance Testing

- **How does this fit into the layered model?**
- **Compliance testing is split out into sections:**

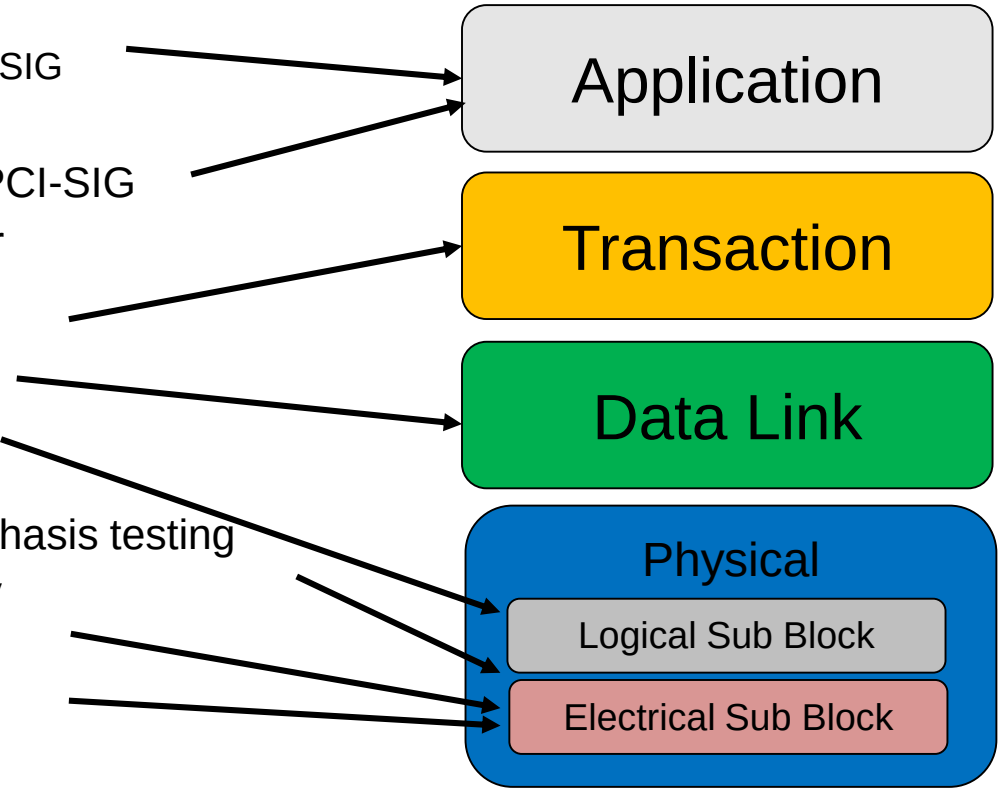


- Configuration Space
  - Using PCIECV tool from PCI-SIG
- System BIOS Testing
  - Using PCIEPT tool from PCI-SIG



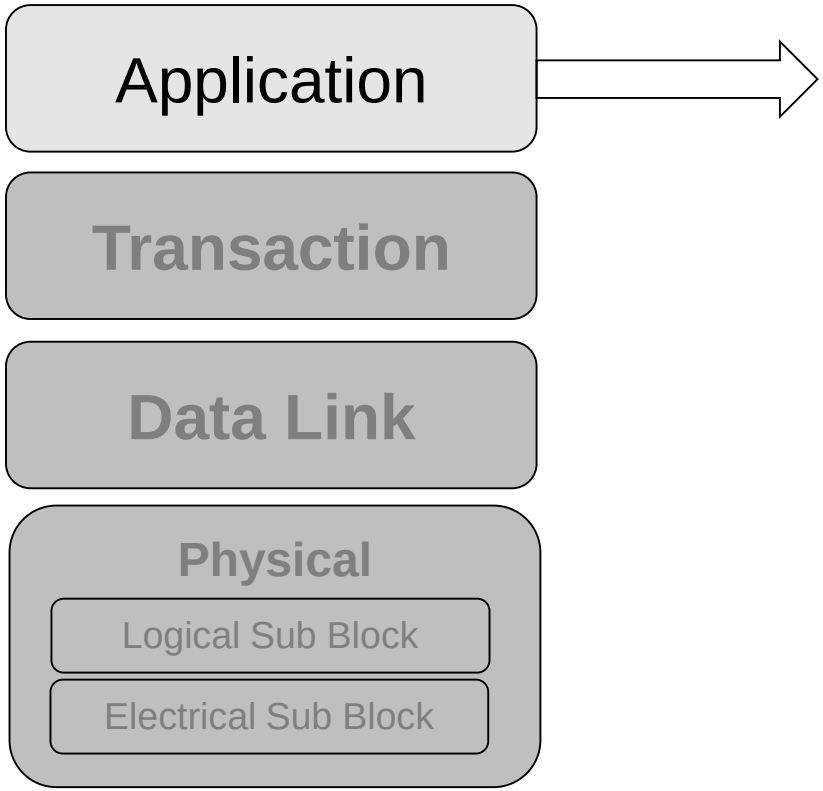
- Link and Transaction Layer
  - Transaction Layer testing
  - Data Link Layer testing
  - Testing Link Training
- Electrical
  - Link Equalization/De-emphasis testing
  - Transmitter Signal Quality
  - Receiver Jitter Tolerance





Copyright © 2017 PCI-SIG® - All Rights Reserved

PCI Express Storage



Software
– Ethernet, NVMe, SOP, AHCI, SATA…

Copyright © 2017 PCI-SIG® - All Rights Reserved

PCIe Storage Devices

- **Where does this fit into the layered model**
  - NVMe
  - PQI/SOP
  - AHCI/SATA Express
- **How do I debug?**
- **“I don’t care about PCI Express!!!”**
- **Compliance Programs for Storage are outside of the scope of PCI-SIG; however, PCI-SIG is impacted by form factor changes**

Copyright © 2017 PCI-SIG® - All Rights Reserved

Device ID Changes

- **Something to watch out for when using the some drivers is that the Bus:Device:Function of the device may change during initialization**
- **It is important when testing NVMe devices to ensure that you are observing the correct Device ID**
- **This also applies to a link upstream of a switch, there may be multiple streams of traffic on one link.**

Copyright © 2017 PCI-SIG® - All Rights Reserved

How Do I Debug These Products?

- **It may take a different way of thinking to use a PCI Express analyzer for debugging a storage problem**
- **The storage protocols reside within the PCI Express payload data, and requires further decoding**
- **It is possible that the only interface to these devices is PCI Express**
- **New form factors are emerging also, U.2, SFF-8639, M.2 for example**

Copyright © 2017 PCI-SIG® - All Rights Reserved

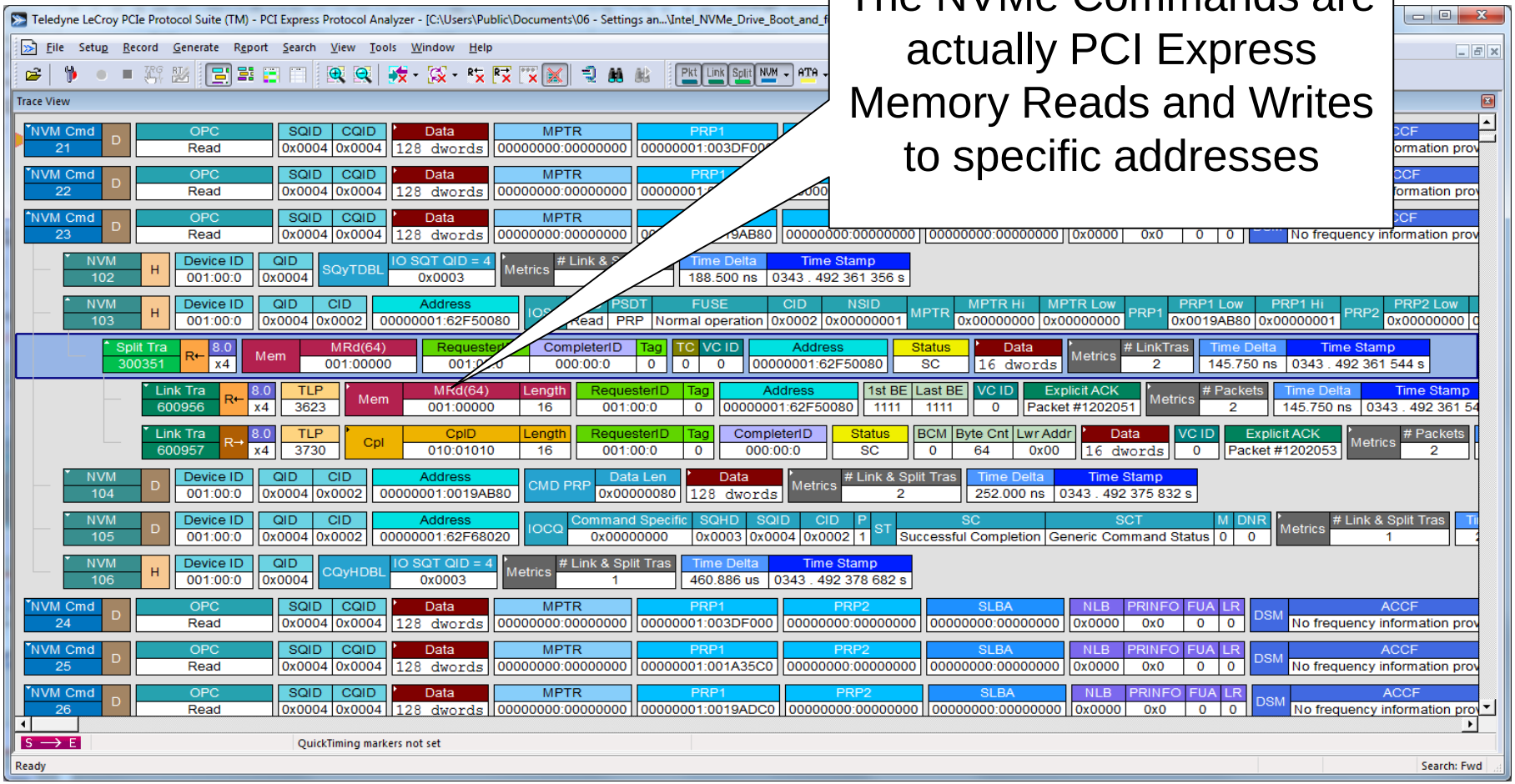
PCIe Storage

- **The underlying PCIe traffic has an impact on performance of the Storage layers**
- **Queue balancing**
- **Possibility of multiple devices over one link**
  - Devices behind a switch
  - Virtualized devices
  - Built in scalability of technologies such as NVMe

Copyright © 2017 PCI-SIG® - All Rights Reserved

NVMe Example

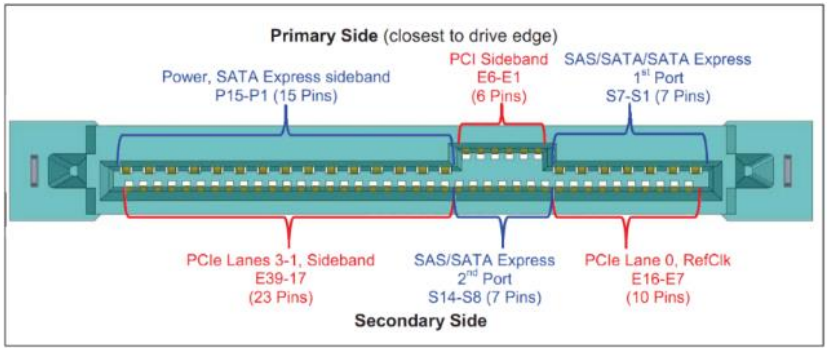
The NVMe Commands are actually PCI Express Memory Reads and Writes to specific addresses



Copyright © 2017 PCI-SIG® - All Rights Reserved

U.2 Connector

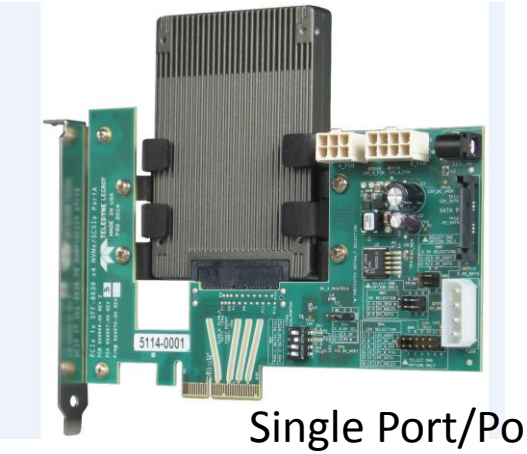
- **Name changed from SFF-8639 to U.2**
- **Connector supports PCIe, SAS, SATA**
- **Supports single or dual link configuration**
  - Dual Port Enable pin grounded for slot that supports dual link
  - Single x4 link or dual x2 links



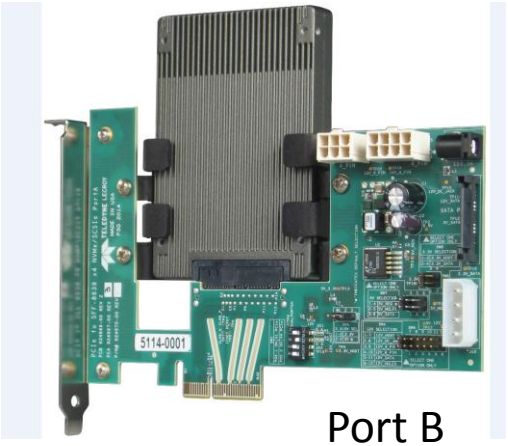
Copyright © 2017 PCI-SIG® - All Rights Reserved

Dual Port U.2

- **Dual Port PCIe/NVMe devices are starting to be developed**
  - A Dual Port device has 2 separate independent PCIe links
  - It requires 2 analyzers and a separate Interposer to debug
- **Compliance Testing requires each port to be tested separately**



Single Port/Port A

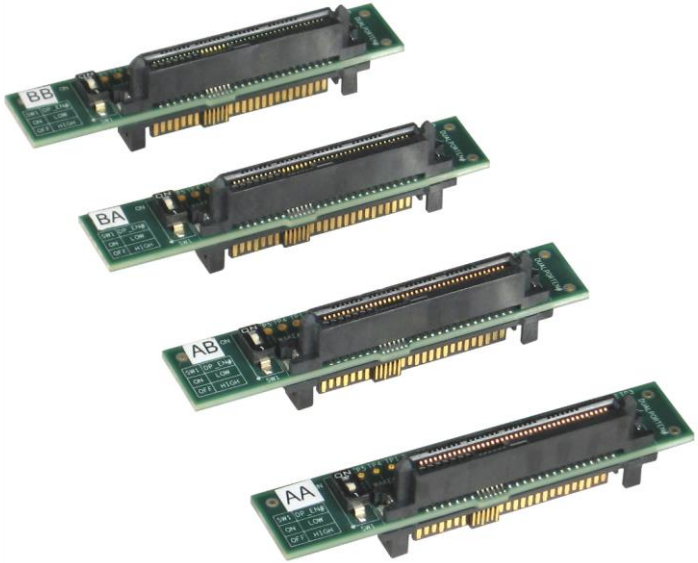


Port B

Copyright © 2017 PCI-SIG® - All Rights Reserved

Debugging Dual Port

- **Try to set up analysis tools separately for each port before trying to look simultaneously**
- **If necessary, use a transposer to block one or other of the ports to work on a single port at a time**



Copyright © 2017 PCI-SIG® - All Rights Reserved

Summary and Conclusions

- **The easiest way to debug PCI Express problems is to determine which layer is showing the problem**
- **Using the correct tool will make the job much easier**
- **Use a “Snapshot” trigger first of all to assess what is going on**
- **Ensure lower level (electrical) layers are compliant and within spec before debugging protocol problems**
- **Remember the PCI-SIG compliance testing is not a substitute for thorough validation, it is a useful sanity check**

Copyright © 2017 PCI-SIG® - All Rights Reserved

Thank you for attending the PCI-SIG Developers Conference 2017.

For more information please go to www.pcisig.com

Copyright © 2017 PCI-SIG® - All Rights Reserved